Big Data Coursework for Computational Medicine

The Mayo Clinic – Rochester

PI: Jyotishman Pathak, Christopher Chute, and Claudia Neuhauser Grant Number: 1R25EB020381-01

As the era of "Big Data" is dawning on biomedical research, multiple types of biomedical data, including phenotypic, molecular (including -omics), clinical, imaging, behavioral, and environmental data is being generated on an unprecedented scale with high volume, variety and velocity. These datasets are increasingly large and complex, challenging our current abilities for data representation, integration and analysis for improving outcomes and reducing healthcare costs. It is well-recognized that the greatest challenge to leveraging the significant potentials of Big Data is in educating and recruiting future computational and data scientists who have the background, training and experience to master fundamental opportunities in biomedical sciences. This demands interdisciplinary education and hands-on practicum training on understanding the application, analysis, limitations, and value of the Big Data. To bridge this knowledge gap for the U.S. biomedical workforce, we propose to develop a research educational program-Big Data Coursework for Computational Medicine (BDC4CM)-that will instruct students, fellows and scientists in the use of specific new methods and tools fo Big Data by providing tailored, in-depth instruction, hands-on laboratory modules, and case studies on Big Data access, integration, processing and analysis. Offered by highly interdisciplinary and experienced faculty from Mayo Clinic and the University of Minnesota, this program will provide a short- term training opportunity on Big Data methods and approaches for: 1) data and knowledge representation standards; 2) information extraction and natural language processing; 3) visualization analytics; 4) data mining and predictive modeling; 5) privacy and ethics; and 6) applications in comparative effectiveness research and population health research and improvement. Our primary educational goal is to prepare the next generation of innovators and visionaries in the emerging, multidimensional field of Big Data Science in healthcare, as well as to develop a future workforce that fulfills industry needs and increases U.S. competitiveness in healthcare technologies and applications. PUBLIC HEALTH RELEVANCE: The postdoctoral Big Data Coursework for Computational Medicine (BDC4CM) program seeks to provide short-term education and hands-on practicum training in utilization of biomedical Big Data. BDC4CM will address a major need for the U.S. biomedical workforce to develop and enhance existing skills in application, analysis, limitations, and value of the Big Data.